Amendments to the Claims

Claim 1 (currently amended). A system for disambiguating speech input using one of voice mode interaction, visual mode interaction, or a combination of voice mode and visual mode interaction with an application comprising:

a speech disambiguation mechanism;

an options and parameters component for receiving user parameters and application parameters for controlling the speech disambiguation mechanism, wherein both the user and the application can set the parameters to control said mechanism, and wherein the parameters include confidence thresholds governing unambiguous recognition and close matches;

a speech recognition component that receives recorded audio, speech or a combination of audio and speech input and generates:

one or more tokens corresponding to the speech input without translation of the language in which the audio, speech, or a combination of speech and audio input is received ; and

for each of the one or more tokens, a confidence value indicative of the likelihood that a given token correctly represents the speech input;

a selection component that identifies, according to a selection algorithm, which two or more tokens generated without translation of the language in which the speech, audio or combination of speech and audio input is received, and presents said tokens as alternatives to the user are to be presented to a user as alternatives, wherein said alternatives are words or tokens;

2

one or more disambiguation components that present the alternatives to the user in one of voice mode, visual mode, or a combination of voice mode and visual mode, and receive an alternative selected by the user from the plurality of alternatives presented to the user in one of voice mode, visual mode, or a combination of voice mode and visual mode; and

an output interface that presents the selected alternative to an application as input.

Claim 2 (cancelled). The system of claim 1, wherein the disambiguation components and the application reside on a single computing device.

Claim 3 (cancelled). The system of claim 1, wherein the disambiguation components and the application reside on separate computing devices.

Claim 4 (original). The system of claim 1, wherein the one or more disambiguation components perform said interaction by presenting the user with alternatives in a visual mode, and by receiving the user's selection in a visual mode.

Claim 5 (original). The system of claim 4, wherein the disambiguation components present the alternatives to the user in a visual form and allow the user to select from among the alternatives using a voice input.

Claim 6 (cancelled). The system of claim1, wherein the one or more disambiguation components perform said interaction by presenting the user with alternatives in a visual mode, and by receiving the user's selection either in a visual mode, a voice mode, or a combination of visual mode and voice mode.

Claim 7 (original). The system of claim 1, wherein the selection component filters the one or more tokens according to a set of parameters.

Claim 8 (original). The system of claim 7, wherein the set of parameters is user specified.

3

Claim 9 (cancelled). The system of claim1, wherein the one or more disambiguation components disambiguates the alternatives in plural iterative stages, whereby the first stage narrows the alternatives to a number of alternatives that is smaller than that initially generated by the selection component, but greater than one, and whereby the one or more disambiguation components operative iteratively to narrow the alternatives in subsequent iterative stages.

Claim 10 (cancelled). The system of claim 9, whereby the number of iterative stages is limited to a specified number.

Claim 11 (currently amended). A method of processing speech input using one of voice mode interaction, visual mode interaction, or a combination of voice mode and visual mode interaction with an application comprising:

a speech disambiguation mechanism;

receiving user parameters and application parameters for controlling the speech disambiguation mechanism, wherein both the user and the application can set the parameters to control said mechanism, and wherein the parameters include confidence thresholds governing unambiguous recognition and close matches;

receiving a speech input from [[a]] the user;

determining whether the speech input is ambiguous;

if the speech input is not ambiguous, communicating a token representative of the speech input to an application as input to the application; and

if the speech input is ambiguous;

4

selecting two or more tokens generated without translation of the language in which the speech, audio or combination of speech and audio input is received, and presenting said tokens as alternatives to be presented to the user as alternatives, wherein said alternatives are words or tokens;

presenting the alternatives to the user in one of voice mode, visual mode, or a combination of voice and visual mode, and receiving a selection of an alternative from by the user from the plurality of alternatives resented to the user in one of voice mode, visual mode, or a combination of voice mode and visual mode; and,

communicating the selected alternative to the application as input to the application.

Claim 12 (original). The method of claim 11, where the interaction comprises the concurrent use of said visual mode and said voice mode.

Claim 13 (original). The method of claim 12, wherein the interaction comprises the user selecting from among the plural alternatives using a combination of speech and visual-based input.

Claim 14 (original). The method of claim 11, wherein the interaction comprises the user selecting from among the plural alternatives using visual input.

5